NTNU Norwegian University of Science and Technology

Exercise No. 2

Course: Numerical Linear Algebra

Submitted By: Muhammad Hamza Khalid

Student Number: 495026

October 21, 2018

Question 5

Original Problem:

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$$

Reformulated Problem:

$$\begin{aligned} x &= A^{-1}\left(b - By\right), \\ x &= (B^T)^{-1}c, \end{aligned} \tag{1}$$

Equating both equations gives,

$$\begin{aligned} (B^T)^{-1}c &= A^{-1}\left(b - By\right), \\ c &= B^T A^{-1} b - B^T A^{-1} B y, \\ \underbrace{B^T A^{-1} B}_{A'} y &= B^T A^{-1} b - c, \end{aligned}$$

The new system can be written as, $A'y = g$, where $g = B^T A^{-1} b - c$.

Properties of the new system:

Matrix A is SPD and B is full rank. A SPD implies that A^{-1} will be SPD.

Symmetric

$$(A')^T = (B^T A^{-1} B)^T = B^T (A^{-1})^T (B^T)^T = B^T A^{-1} B = A$$

Positive Definite

Let $z \in R^n \backslash \{0\}$, A being SPD implies $Q \Lambda Q^T$ such that $QQ^T = I$. Another assumption is that B is full rank. Consider,

$$z^T B^T A^{-1} B z = z^T B^T Q^{-1} \Lambda^{-1} Q B z = (QBz)^T \Lambda^{-1} (QBz)$$

Since A is SPD and the eigen values are greater than zero, i.e $\lambda_i > 0$, also A^{-1} will be SPD and eigen values of $A^{-1} = \frac{1}{\lambda_i} > 0$ for $i = 1 : n$. Using the result from Rayleigh quotient,

$$R(QBz) = \tfrac{(QBz)^T \Lambda^{-1} (QBz)}{(QBz)^T (QBz)} = \tfrac{(QBz)^T \Lambda^{-1} (QBz)}{(zB)^T (zB)}$$

Using the result of Rayleigh quotient we get, $R(QBz) \in [\lambda_1, \lambda_n]$ and since $\lambda_i > 0$ for all i, hence we conclude our new system will be **positive definite**.

Hence the reformulated system is Symmetric Positive Definite and the iterative method used to solve the problem was **Nested Conjugate Gradient method**.

The **condition number** of the original problem was 3.9×10^3, and reformulating the problem greatly reduce the condition number to 256.4. Time taken by the algorithm was 0.22 seconds and number of iterations were 28. An error plot for tolerance 10^{-6} is shown in Figure 1.

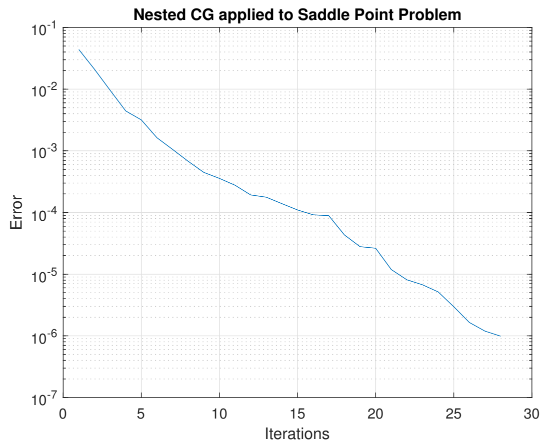

Figure 1: Iterations vs Error

Question 6

6.1

The matrix $A = (L + \delta x^2 k^2 I)$.In order to check positive definiteness we look for the upper left determinants of the matrix. The first upper left determinant tells, $k^2 > \frac{4}{dx^2}$ whereas the second upper left determinant tells, $k^2 > \frac{4}{dx^2}$. Following the same trend we end up with a condition that, for A to be SPD $k^2 > \frac{8}{dx^2}$. If this condition is satisfied than our matrix A is guaranteed SPD.

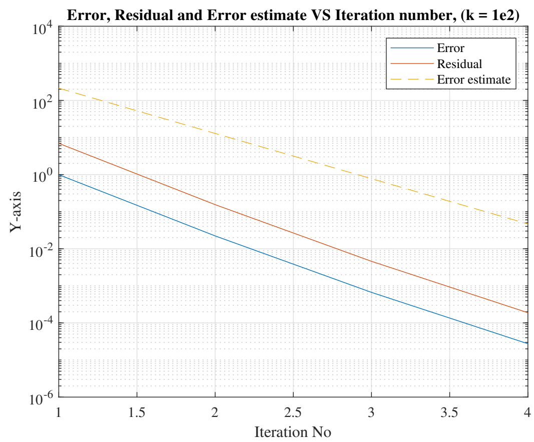

Figure 2: Comparison of error, residual and error estimate CG, k big

Figure 2 is the plot of residual, error and error estimate of Conjugate gradient right hand side, with k large. However if we increase the value of k the error is still decreasing but after 4 iterations the convergence slope starts to bend.

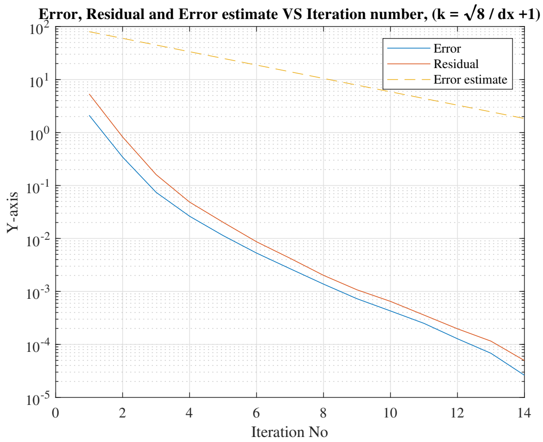

Figure 3: Comparison of error, residual and error estimate CG, k small

6.2: Performance of different Pre-conditioners

Convergence of different Pre-conditioners

Figure 4 shows the convergence of different preconditioners. It is interesting to note that changing the values of k influence the speed of convergence for different preconditioners, which will be discussed in subsection discussing how k affect convergence. Figure 4 is the plot with a small value of k. It is interesting to note that, symmetric GS converged in less iterations as compared to the fast poisson solver.

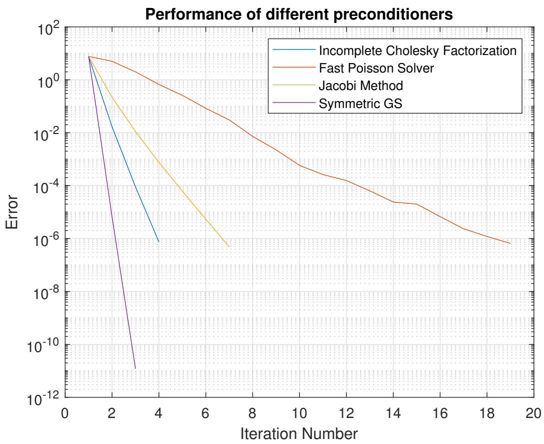

Figure 4: Convergence of different Pre-conditioners

Time

Table 1 shows the time taken by the different pre-conditioners for solving the problem. Difference in time was seen for each preconditioner depending on the k. For small k Fast Poisson solver took the most time as compared to other preconditioners. This time can be improved by the use of discrete sine transform which will also improve the memory requirement of the problem.

Preconditioner	Avg. Time (k small)	Avg. Time (k large)
Incomplete Cholesky Factorization	0.003895	0.004786
Fast Poisson Solver	0.022846	0.013422
Jacobi Method	0.008255	0.001411
Symmetric GS	0.049780	0.050148

Table 1: Average Time

Error estimates

The problem regarding the evaluation of spectral radius was unclear, so I decided to use the error estimate for conjugate gradient method for the new system i.e after preconditioner was applied and plot them. For a large value of k results observed can be seen in the Figure 5.

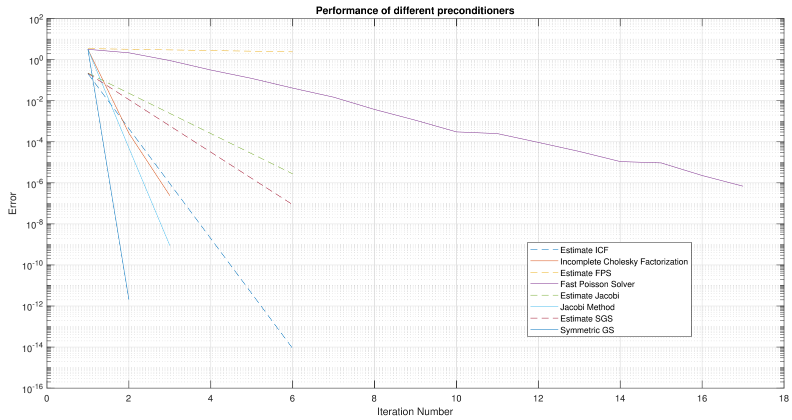

Figure 5: Convergence and error estimate of different Pre-conditioners

Effect of k and Δx on convergence

Like said in the previous section it was not completely understood how to find the matrix A1 and A2 in order to find the iteration matrix to check convergence. However, during the experiments it was observed that, increasing the value of k helped in fast convergence of the above method.

Theoretically it can be explained using the fact that, a high value of k results in diagonal dominant matrix, which results in good convergence as per lemma in Y.Saad's book.

P#1 ⓐ Derive the basic version of GMRES by using the standard formula,

$$\tilde{x} = x_0 + V(W^TAV)^{-1}W^T r_0 \quad \text{with} \quad V = V_m, \; W = AV_m$$

Sol: let $V = [v_1, \ldots v_m]$ an $n \times m$ matrix with coulmn vector as basis of K.

and $W = AV = AV_m$

Suppose the approximate sol. is given by,

$$\tilde{x} = x_0 + V(W^TAV)^{-1}W^T r_0$$

then minimising the residual for GMRES can be written as,

$$R(\tilde{x}) = \min_{x \in x_0 + K} \|b - Ax\|_2 \longrightarrow ①$$

using the fact $\tilde{x}$ an approximate sol.

Consider, $b - A\tilde{x} = b - A(x_0 + V(W^TAV)^{-1}W^T r_0)$

$$= \underbrace{b - Ax_0}_{r_0} - AV_m \underbrace{(W^TAV_m)^{-1}W^T r_0}_{y_m}$$

$$= r_0 - AV_m y_m$$

"Here y_m comes from the orthogonality condition which leads to the system $W^TAVy = W^T r_0$"

using "$AV_m = V_{m+1}\bar{H}_m$"

$$\Rightarrow \quad = \beta v_1 - V_{m+1}\bar{H}_m y_m$$

$$= V_{m+1}(\beta e_1 - \bar{H}_m y_m)$$

Since the coulmn vectors of V_{m+1} orthonormal, then

$$① \Rightarrow R(y) = \min_{x \in x_0 + K} \|b - Ax\|_2$$

Suppose $\tilde{x}$ minimize the above expression,

$$\Rightarrow R(y) = \|\beta e_1 - \bar{H}_m y_m\|_2 \longrightarrow ②$$

Hence we conclude that a GMRES approximation is unique vector of $x_0 + K$ that minimizes ②

The approximation can be obtained using

$$\tilde{x} = x_0 + V(W^TAV)^{-1}W^T r_0$$

$\boxed{\text{VM}}$

b) Consider the solution of the linear system $Ax=b$

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

i) Compute the matrices V_m and H_m for $m=1,\dots 5$ resulting from the application of arnoldi Algorithm.

Sol: For our particular problem

$$r_0 = b - Ax_0$$

$$\Rightarrow \boxed{r_0 = b}$$

Choosing, $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e_1$

and applying the Arnoldi algorithm we get orthonormal basis of Krylov subspace which are infact canonical basis of $\mathbb{R}^5$.

for $m=0,1\dots 4$

$$A^m r_0 = A^m b = A^m e_1$$

i.e if we write $K_m(A, r_0) = \{r_0, Ar_0, A^2r_0, A^3r_0, A^4r_0\}$

$$\Rightarrow K_m(A,b) = \{b, Ab, A^2b, A^3b, A^4b\}$$

in our case $b = e_1$

$$\Rightarrow K_m(A,b) = \{e_1, Ae_1, A^2e_1, A^3e_1, A^4e_1\}$$

$$\boxed{V_m = \{e_1, e_2, e_3, e_4, e_5\}}$$

After m-steps of Arnoldi iteration, $V_m = [e_1, e_2 \cdots e_5]$

and H_m by,

$$\Rightarrow V_m^T A V_m = H_m$$

Since $V_m = I_{5\times 5}$

$$\Rightarrow \boxed{H_m = A}$$

Arnoldi algorith breaks down at $h_{6,5}$ i.e $h_{6,5} = 0$.

[ii] Compute the FOM iterates y_m, x_m. (when possible)

Sol: Since $K_m(A, r_0) = \{r_0, Ar_0, \cdots A^{m-1} r_0\}$

in our case $m = 1, \cdots 5$

$$\Rightarrow K_5(A, r_0) = \{r_0, Ar_0, A^2 r_0, A^3 r_0, A^4 r_0\}$$

For $m = 1:4$ The matrices H_m are singular with the first row identical to zero.

Therefore it is not possible to solve

$$y_m = H_m^{-1}(Be_1).$$

However for $m = 5$, $H_5 = A$ and $V_5 = I_{(5\times 5)}$

$$\Rightarrow y_5 = H_5^{-1}(Be_1) \text{ where } B = \|r_0\|_2 = \|b\|_2$$

i.e $B = \|e_1\|_2 = 1$

$$\Rightarrow y_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = e_5$$

Now the iterate $x_5 = x_0 + V_5 y_5 = e_5$

Therefore $\boxed{y_5 = e_5}$ and $\boxed{x_5 = e_5}$

[iii] Describe in detail the QR factorization of the matrix $\tilde{H}_m$, for $m=1,\cdots 5$ using given rotations.

Sol: As computed in the previous problem Arnoldi algorithm breaks down at $h_{6,5}$ i.e $h_{6,5}=0$

Therefore we can write $\tilde{H}_5$ as,

$$\tilde{H}_5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The next step involves finding the rotations defined by,

$$\Omega_i = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c_i & s_i & \\ & & & -s_i & c_i & \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} \quad \text{s.t } c_i^2+s_i^2=1$$

For the first iteration

$$\Omega_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{as,} \quad \boxed{\begin{aligned} c_1 &= \frac{h_{11}}{\sqrt{h_{11}^2+h_{21}^2}} = 0 \\ s_1 &= \frac{h_{21}}{\sqrt{h_{11}^2+h_{21}^2}} = 1 \end{aligned}}$$

Likewise $c_2=0$, $s_2=0$

After looking at the structure of $\tilde{H}_5$ it is easy to deduce that for $k=2,3,4$ $\Omega_k=\Omega_1$.

From the theory we can write Q_5 as,

$$Q_5 = \Omega_5\Omega_4\Omega_3\Omega_2\Omega_1$$

i.e the product of Ω_i. Our aim is to construct an upper triangular matrix and since it is already an upper triangular matrix, therefore we can take $Q_5 = I_{2\times 2}$.

Since,

$$\bar{R}_m = Q_m \bar{H}_m$$

$$\bar{R}_5 = \begin{pmatrix} Q_4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{H}_4 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\bar{R}_5 = \begin{pmatrix} Q_4 \bar{H}_4 & Q_4 e_1 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \tilde{R}_4 & e_5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Q_4 \tilde{H}_4 & Q_4 e_1 \\ 0 & 0 \end{pmatrix}$$

[iv] Compute the GMRES iterates x_m and y_m

Sol: Since Q_m is unitary

$$\Rightarrow \min \|\beta e_1 - \bar{H}_m y\|_2 = \min \|\bar{g}_m - \bar{R}_m y\|_2$$

where $\bar{g}_m = Q_m(\beta e_1)$, with $\beta = \|r_0\|_2 = 1$

we have, $\bar{g}_m = (-1)^m e_{m+1}$ for $m = 1:4$

we have $y_m = R_m^{-1} \bar{g}_m$ where $R_m = I_{m\times m}$

using both we get $\boxed{y_m = 0}$ for $m = 1:4$ and $\boxed{y_5 = e_5}$

$\Rightarrow$ $\boxed{x_m = x_0}$ for $m = 1:4$ and $\boxed{x_5 = e_5}$

Problem #2

(i) If $A \in \mathbb{R}^{n \times n}$ is symmetric, then A is positive definite if following conditions are satisfied

(a) $x^T A x > 0 \quad \forall\, x \in \mathbb{R}^n \quad x \neq 0$

(b) All eigen values are positive.

Show (a) and (b) are equivalent.

Sol: let $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix

(a) => (b) :- (a) implies that $x^T A x > 0$

let λ be the eigenvalue and $\vec{x}$ be the corresponding eigen vector, then,

$$A\vec{x} = \lambda \vec{x} \qquad \forall\, x \neq 0$$

Multiply both sides with "x^T"

$$\vec{x}^T A \vec{x} = \vec{x}^T \lambda \vec{x}$$

$$\Rightarrow \vec{x}^T \lambda \vec{x} = \vec{x}^T A \vec{x} > 0$$

$$\Rightarrow \lambda \vec{x}^T \vec{x} = \vec{x}^T A \vec{x} > 0$$

$$\Rightarrow \lambda \|\vec{x}\|^2 > 0$$

This will be positive due to square of norm.

in order for the above expression to be positive λ must be greater than zero.

Therefore, we say that all eigen values corresponding to the eigen vector should be greater than zero.

i.e $\lambda_i > 0 \qquad i = 1, \ldots n$

Now, (b) => (a) (b) => All eigenvalues are greater than 0.

Since every real symmetric matrix is diagonalizeable,

$\Rightarrow \quad A = QDQ^T, \quad Q$: orthogonal matrix

D: Diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Here λ_i are the eigen values and they are greater than zero from the assumption.

let $\vec{x} \neq 0 \in \mathbb{R}^n$

$$\vec{x}^T A \vec{x} = \vec{x}^T Q D Q^T \vec{x}$$

Let $\vec{y} = Q^T \vec{x}$

$$\Rightarrow \vec{x}^T A \vec{x} = \vec{y}^T D \vec{y}$$

Consider the R.H.S of the above equation

$$\Rightarrow \vec{y}^T D \vec{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\vec{y}^T D \vec{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2 > 0$$

$\boxed{\text{Since } \lambda_i > 0}$

$$\Rightarrow \vec{x}^T A \vec{x} = \vec{y}^T D \vec{y} > 0$$

$$\Rightarrow \vec{x}^T A \vec{x} > 0$$

Therefore we conclude that (a) and (b) conditions are equivalent.

(b) Show that the following equivalence.

$$A \text{ is SPD} \Leftrightarrow A^{-1} \text{ is SPD}$$

Sol: $\Rightarrow$ Assume $A \in \mathbb{R}^{n\times n}$ is SPD

Consider the inverse of A i.e A^{-1}

Aim: to show A^{-1} is SPD

Since A is invertible we can write

$$(A^{-1})^T = (A^T)^{-1}$$

A being symmetric $A^T = A$

$$(A^{-1})^T = (A^T)^{-1} = A^{-1}$$

$$\Rightarrow (A^{-1})^T = A^{-1} \Rightarrow A^{-1} \text{ is symmetric}$$

To show positive definess,

Since A is SPD

$$\Rightarrow A = QDQ^T$$

$$A^{-1} = (QDQ^T)^{-1}$$

$$A^{-1} = (Q^T)^{-1} D^{-1} Q^{-1}$$

we know that $Q^{-1} = Q^T$

$$A^{-1} = QD^{-1}Q^T$$

As, $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ so, $D^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_n} \end{bmatrix}$

Since $\lambda_i > 0 \Rightarrow \frac{1}{\lambda_i} > 0$

Now for $\vec{x} \in \mathbb{R}^{n\times n} \neq 0$

$$\vec{x}^T A^{-1} \vec{x} = \vec{x}^T Q D^{-1} Q^T x = y^T D^{-1} y > 0$$

Since,
$$\vec{x}^T A^{-1} \vec{x} > 0$$
$\Rightarrow$ A^{-1} is Positive definite

$\Leftarrow$ Assume A^{-1} is SPD

To Show: A is SPD

Symmetric: Consider, $AA^{-1} = I_{n\times n}$

Taking transpose
$$(AA^{-1})^T = I^T$$
$$\underbrace{(A^{-1})^T}_{A^{-1} \text{ is SPD} \Rightarrow (A^{-1})^T = A^{-1}} A^T = I$$
$$A^{-1}A^T = I$$
$A^T = A$ $\Rightarrow$ A is symmetric.

Positive Def:

ince A^{-1} is positive definite

$\Rightarrow$ $\vec{x}^T A^{-1} \vec{x} > 0$ for $\vec{x} \in \mathbb{R}^{n\times n} \neq 0$

$A \in \mathbb{R}^{n\times n}$ is symmetric as shown before.

$\Rightarrow$ $A = QDQ^T$ where $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

From the positive defitness of A^{-1}, $\frac{1}{\lambda_i} > 0$

Therefore we say $\lambda_i > 0$ ie all eigen values are greater than 0.

$\Rightarrow$ A is positive definite
$\Rightarrow$ A is SPD.

Hence we conclude,
$$\boxed{A \text{ SPD} \Longleftrightarrow A^{-1} \text{ SPD}}$$

$\boxed{iii}$ If $A \in \mathbb{R}^{n \times n}$, the quotient

$$R(x) = \frac{x^T A x}{x^T x}$$

is known as Rayleigh quotient of A.

Show that if A is positive def., the A satisfies

$$\lambda_n \leq R(x) \leq \lambda_1 \quad \forall\, x \in \mathbb{R}^n,\ x \neq 0$$

Sol:- Assume A to be SPD.

$$\Rightarrow \quad x^T A x > 0$$

and $\lambda_i > 0$ for $i = 1, \dots n$

$$R_A(x) = \frac{x^T A x}{x^T x} \longrightarrow (i)$$

A being symmetric we can diagonalize it as,

$$A = QDQ^T$$

$$(i) \Rightarrow \quad R_A(x) = \frac{x^T Q D Q^T x}{x^T x} = \frac{(Qx)^T D (Qx)}{x^T x} \longrightarrow (ii)$$

Consider $(Qx)^T (Qx) = x^T \underbrace{Q^T Q}_{= I} x = x^T x$

$$(ii) \Rightarrow \quad R_A(x) = \frac{(Qx)^T D (Qx)}{(Qx)^T (Qx)} = \frac{y^T D y}{y^T y} = R_A(y)$$

$$R_A(x) = R_A(y) = \frac{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots \lambda_n y_n^2}{y_1^2 + y_2^2 + \dots + y_n^2}$$

A being Symmetric, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \underbrace{> 0}_{\text{(Positive Def)}}$

for a vector $y_1 = (1, 0, 0 \dots 0)$ and $y_n = (0, \dots 0, 1)$

$R_A(y_1) = \lambda_1$ and $R_A(y_n) = \lambda_n$

Since $\lambda_1 \geq \lambda_n$ as λ_1 being the max eigen Value

$$\Rightarrow \quad \lambda_n \leq R_A(x) \leq \lambda_1 \quad \forall\ x \in \mathbb{R}^n / \{0\}$$

Problem 3: let $v_1, v_2, \ldots v_m$ be vectors generated by Arnoldi algorithm. let $P_i = v_i v_i^T$ be orthogonal projection onto span$\{v_i\}$. show,

$$(I-P_j)(I-P_{j-1})\cdots(I-P_1) = I - P_j - P_{j-1} - \cdots - P_1$$

Show that iteration j with classic Gram-Schmidth corresponds to

$$w_j = (I - P_j - P_{j-1} - \cdots - P_1) A v_j$$

while modified Gram Schmidth,

$$w_j = (I-P_j)(I-P_{j-1})\cdots(I-P_1) A v_j$$

Sol:

Consider, $(I-P_j)(I-P_{j-1})\cdots(I-P_1)$

$$= (I - P_j - P_{j-1} + \underbrace{P_j P_{j-1}}_{= \underbrace{P_j v_j}_{=0} v_j^T}) (I - P_{j-2}) \cdots (I-P_1)$$

(Due to orthognality)

$$\Rightarrow (I - P_j - P_{j-1})(I - P_{j-2})\cdots(I-P_1)$$

$$= (I - P_j - P_{j-1} - P_{j-2} + \underbrace{P_j P_{j-2}}_{=0} + \underbrace{P_{j-1} P_{j-2}}_{=0})(I - P_{j-3}) \cdots (I-P_1)$$

$$= (I - P_j - P_{j-1} - P_{j-2})(I - P_{j-3}) \cdots (I - P_1)$$

$\vdots$ (Following the trend)

$$= I - P_j - P_{j-1} - P_{j-2} - P_{j-3} - \cdots P_1$$

$$\Rightarrow (I-P_j)(I-P_{j-1})\cdots(I-P_1) = I - P_j - P_{j-1} - \cdots P_1$$

using classic Gram Schmidth

at iteration j

$$h_{ij} = (Av_j, v_i) = v_i^T A v_j$$

$$w_j = Av_j - \sum_{i=1}^{j} h_{ij} v_i = Av_j - h_{1j} v_1 - h_{2j} v_2 - \cdots h_{jj} v_j$$

$$w_j = Av_j - v_1^T A v_j v_1 - v_2^T A v_j v_2 - \cdots - v_j^T A v_j v_j$$

$$w_j = (I - v_1^T v_1 - v_2^T v_2 - \cdots - v_j^T v_j) A v_j$$

$$\Rightarrow \quad w_j = (I - P_j - P_{j-1} - \cdots P_1) A v_j$$

Now for the Modified Gram Schmidth

At the j-th iteration

$$w_j = A v_j$$

for $i = 1:j$

$$h_{ij} = (w_j, v_i) = (A v_j, v_i) = v_i^T A v_j$$

$$w_j = w_j - h_{ij} v_i$$

end

Subsituting values of w_j and h_{ij} and iterating from $i=1:j$ will give the following expression

$$w_j = A v_j - (v_i^T A v_j) v_i$$

$$w_j = (I - v_i^T v_i) A v_j$$

$$w_j = (I - v_1^T v_1)(I - v_2^T v_2) \cdots (I - v_j^T v_j) A v_j$$

$$w_j = (I - v_1 v_1^T)^T (I - v_2 v_2^T)^T \cdots (I - v_j v_j^T)^T A v_j$$

using $(AB)^T = B^T A^T$

$$w_j = ((I - P_j)(I - P_{j-1}) \cdots (I - P_1))^T A v_j$$

$$\Rightarrow \quad w_j = (I - P_j)(I - P_{j-1}) \cdots (I - P_1) A v_j$$

Problem 4:- We consider the Saddle point problem,

$$\begin{pmatrix} A & D \\ B^T & O \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$$

$A \in \mathbb{R}^{n\times n}$
$B, D \in \mathbb{R}^{n\times k}$ $\quad O \in \mathbb{R}^{k\times k}$, $k \le n$.
Assume A is diagonalizable i.e $A = X^{-1}\Lambda X$
B and D s.t rank(B) = rank$(D) = k$. and they are factorized as,

$$B = X^{T}\Gamma X \qquad D = X^{-1}\Gamma Q$$

$\Gamma \in \mathbb{R}^{n\times k}$ – Diagonal matrix with elements $0.5 \le \gamma_{ii}^{2} \le 1$ $\quad i=1,\dots k$
Q – orthogonal matrix i.e $Q^TQ = I$

(a) Eliminate x from the system and find

$$Q^T\Gamma^T Q\Lambda^{-1}Qy = g \quad , \quad g = B^TA^{-1}b - c$$

Sol:

System can written in the form of linear eq's as,

$$Ax + Dy = b \longrightarrow (i)$$
$$B^Tx + Oy = c \longrightarrow (ii)$$

$$(i) \Rightarrow \quad Ax = b - Dy$$
$$x = A^{-1}(b - Dy)$$

$$(ii) \Rightarrow \quad B^TA^{-1}(b - Dy) + Oy = c$$
$$B^TA^{-1}b - B^TA^{-1}Dy + Oy = c$$
$$\underbrace{B^TA^{-1}b - c}_{=g} = B^TA^{-1}Dy - Oy$$

$$\Rightarrow \quad B^TA^{-1}Dy - Oy = g$$

Since $\quad A = X^{-1}\Lambda X$
$\quad A^{-1} = X\Lambda^{-1}X^{-1}$

using this in the equation above,

$$\Rightarrow \quad B^TX\Lambda^{-1}X^{-1}Dy - Oy = g$$

$$B^T X \Lambda^{-1} X^{-1} D y - O_y = g \longrightarrow (*)$$

Now using
$$B = X^T \Gamma Q$$
(taking transpose) $\Rightarrow$ $B^T = Q^T \Gamma^T X$

Subsituting value of B^T and D

$$\Rightarrow (Q^T \Gamma^T X)(X \Lambda^{-1} X^{-1})(X^{-1} \Gamma Q) y = g$$

$$\Rightarrow \boxed{Q^T \Gamma^T \Lambda^{-1} \Gamma Q y = g}$$

"Reformulated system"

ⓑ Assume eigen values of A^{-1},

$$1 \leq \mathbb{R}(\lambda_n(A^{-1})) \leq \mathbb{R}(\lambda_{n-1}(A^{-1})) \cdots \leq 4$$

and $\quad 0 \leq |\tau(\lambda_i)(A^{-1})| < 0.25$

for $m=2$ find upper bound for

$$\frac{\|r_m\|_2}{\|r_0\|_2}$$

Sol: We are aiming to solve the problem,

$$Q^T \Gamma^T \Lambda^{-1} \Gamma Q y = \underbrace{B^T A^{-1} b - c}_{=g}$$

or $$\underbrace{(\Gamma Q)^T \Lambda^{-1} (\Gamma Q)}_{= A1} y = g$$

$A1$ is symmetric as, $A1^T = Q^T \Gamma^T \Lambda^{-1} \Gamma Q = (\Gamma Q)^T \Lambda^{-1} (\Gamma Q) = A1$
also, $A1$ diagonalizeable.

Eigen values of $A1$ are similar to A^{-}.

Using Proposition 6.32 from Y. Saad book

$$\frac{\|r_m\|_2}{\|r_0\|_2} \leq K_2(X) \min_{\substack{p \in \mathbb{P}_m \\ p(0)=1}} \max_{i=1:n} |P_m(\lambda_i)|$$

Replacing the minimal polynomial by

$$P_m(\lambda) = \frac{C_m(t(\lambda))}{C_m(t(0))}$$

and using theorem 6.29 to estimate

$$\frac{\|r_m\|_2}{\|r_0\|_2} \leq 2 K_2(X) \left(\frac{\sqrt{K(A1)} - 1}{\sqrt{K(A1)} + 1} \right)^m$$

$K_2(X) \approx 1$ Since A is normal.

Now for $m=2$,

$$\frac{\|r_2\|_2}{\|r_0\|_2} \leq 2 \left(\frac{\sqrt{K(A1)} - 1}{\sqrt{K(A1)} + 1} \right)^2 \longrightarrow 1$$

Since $K(A1) = \max_{i=1:n} |\lambda_i|$

from the given data, $K(A) = 4$

Subsituting value in 1

$$\Rightarrow \frac{\|r_2\|_2}{\|r_0\|_2} \leq 2 \left(\frac{2-1}{2+1} \right)^2$$

$$\leq 2 \left(\frac{1}{3} \right)^2 = \frac{2}{9}$$

$$\Rightarrow \boxed{\frac{\|r_2\|_2}{\|r_0\|_2} \leq \frac{2}{9}}$$